DOCUMENT RESUME

ED 193 259 TM 800 573

AUTHOR Gould, Jewell; Ward, James
TITLE Teachers and Tests: The Teacher Union Response.
PUB DATE Jul 80
NOTE 17p.; Paper presented at the Annual Summer Instructional Leadership Conference of the American Association of School Administrators (Chicago, IL, July, 1980). For related document see TM 800 572.

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Educational Improvement; *Educational Testing; Elementary Secondary Education; National Surveys; Public Education; *Standardized Tests; *Teacher Attitudes; Testing Problems; *Testing Programs; Unions
IDENTIFIERS *American Federation of Teachers; Test Use

ABSTRACT

Members (elementary and secondary teachers) of the American Federation of Teachers (AFT) were surveyed in preparation for a series of workshops to improve teachers' use and understanding of tests. Results showed generally positive attitudes towards tests, and a specific interest in using tests for diagnostic and prescriptive planning purposes. Standardized tests were seen as least useful for course content modification and instructional program modification. A negative side appeared also: teachers idenfified a variety of factors inhibiting their use of standardized tests. These include: (1) results do not provide enough prescriptive information; (2) results do not provide adequate profiles of student strengths/weaknesses; (3) results are returned too late (or not at all) to be useful; (4) test content does not match curriculum; and (5) test materials are inappropriate for students. Results also indicated, however, that teachers desire maintenance of quality standards in schools and view competency testing as a useful method. Additional research is discussed, supporting the findings that teachers see tests as reliable sources to be used as part of a total educational program. The AFT resolution on testing, and how the AFT training program is designed are summarized. (GK)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
J. Gould
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U S DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED193259

TM 800573

TEACHERS AND TESTS: THE TEACHER UNION RESPONSE

Jewell Gould and James Ward*
American Federation of Teachers
Washington, D.C.

ABSTRACT

Measurement and assessment inquiry as of late has suffered from excesses at both the technical and simplistic extremes of the science. The American Federation of Teachers believes that teachers, the front line participants in testing, have not had an opportunity to fully participate in the debate. As a result, many unanticipated consequences have developed as attention is directed away from the primary task of teaching and towards the particular idiosyncrasies of the test instrument. The American Federation of Teachers, working under a grant from the National Institute of Education, surveyed its members in preparation for a series of workshops designed to improve teachers' use and understanding of tests. The survey showed generally positive attitudes toward tests by teachers and a specific interest in the use of tests for diagnostic and prescriptive planning purposes. Additional independent research is discussed which supports the findings that teachers feel tests are useful and can be relied on when used as part of a total educational program. Tests are useful in identifying students' needs and preparing and delivering the educational program designed to meet the identified goals. This orderly process is useful in establishing public school accountability, a necessity if public support is to remain available for schools. A summary of the most recent AFT resolution on testing and how the AFT training program is designed concludes the report.

Paper presented at the annual Summer Instructional Leadership Conference of the American Association of School Administrators, Chicago, Illinois, July 1980

*Assistant Director and Director of AFT Department of Research, respectively

TEACHERS AND TESTS:
THE TEACHER UNION RESPONSE

Jewell Gould and James Ward*
American Federation of Teachers
Washington, D.C.

We in the American Federation of Teachers believe that schools make a difference. We believe that teachers are important. Furthermore, we believe that the more good information that is available for decisions that must be made about students in the educational process, the better off we all are. The recent concern about more interest in the affective goals in education does not, in our mind, diminish at all the cognitive goals of schooling. Schools which do not teach important cognitive knowledges and skills are failures, and we need assessment procedures and instruments to ensure that our schools are fulfilling their right task.

Educational testing should fulfill that need. Current testing practices, however, are not perfect and much debate surrounds testing. Some of the discussion has been useful and some has served to obfuscate the issues. The current controversy over testing regulation, or truth-in-testing as it is popularly termed, has heightened the controversy.

Walt Haney, in an article entitled "Trouble Over Testing," reported one of the ripples in the troubled water when he wrote.

Paper presented at the annual Summer Instructional Leadership Conference of the American Association of School Administrators, Chicago, Illinois, July 1980

*Assistant Director and Director of AFT Department of Research, respectively

The popular prominence of current debate over testing was perhaps epitomized by the appearance of Ralph Nader on the "Tonight Show" on January 24, 1980, to publicize the recent Nader report on the Educational Testing Service (ETS). After condemning the "reign" of ETS, Nader gave an impassioned plea for wider consideration of traits like perseverance, wisdom, idealism, and creativity--traits that cannot be measured by multiple-choice aptitude and achievement tests of the sort ETS publishes. The "Tonight Show" audience broke into spontaneous applause.

The applause is understandable. The Tonight Show audience has been known to break into applause over even stranger acts including Charo, The Zoo Lady and the Great Carnack, mystic of the East. The scene evoked is Johnny Carson holding an SAT to his forehead, Nader looking on in astonishment, and Carson pronouncing the words perseverance, wisdom, idealism, and creativity. A pause, ensues and Ralph blurts out "What important traits are NOT measured by multiple choice aptitude tests?"

So goes much of the inquiry regarding test use these days. Often the literature available is either too esoteric for use by any but a select group of psychometricians or is so simplistic that its usefulness if compromised. That which is not obfuscated with median validity coefficients may just as well have been plucked from hermetically sealed mayonnaise jars of the Tonight Show and held to the forehead of the talk show host for interpretation. The American Federation of Teachers is anxious to see improvements in testing and assessment programs and instruments. Teachers have generally positive attitudes towards tests and use test data for diagnostic purposes. The goals of AFT and the practices of teachers are compatible with this and call for reasoned debate and actions. The consideration of the unanticipated consequences of change is a critical part of the process. It is this lack of consideration of the po-

tential damages that presents the most frightening aspect of the sweeping changes proposed by test critics today. A few illustrations help point out the problems created by these unanticipated consequences.

A recently published review of the Scholastic Aptitude Test (SAT) was prepared by Warner Slack and Douglas Porter (Harvard Educational Review, May 1980), a pair of Harvard Medical School doctors. The article undercuts, possibly unintentionally, the need for standards in education. Too much emphasis is placed on the effects of the SAT for students failing to achieve their goals of entry into the college or university of their first choice. The myriad problems confronting public education today are passed over as possible reasons affecting students' performance. Public schools have been for decades reaching out to the disadvantaged, the non-English speaking, the handicapped, and the non-college bound students of our schools, seeking to provide opportunity and motivation for such students. This phenomenon is downplayed as Slack characterizes the SAT as a measure of "nothing more than learned information that can be readily acquired by students if they have motivation and opportunity" (emphasis added). Porter, echoing Slack, sets us all up quite nicely. He insists that he never believed intellectual equipment was necessary for students "to make a go of college or become doctors or lawyers or whatever they wish." "Given the opportunity and the schooling, almost everyone can attain their educational goals," Porter continued.

They decry the SAT questions containing "little used vocabulary" and "tricky math" that penalizes students from poor families and minority groups, or students with poor preparation in language skills and math. Slack and Porter contend that these students' experiences with language, literature, and math are likely to differ from those of the test designers. Either we are to assume that the

tests are constructed using language that the student may be called upon to use in the pursuit of higher education or are we to assume that everyone, regardless of their skill level, is entitled to pass the SAT. Our assumption is that the SAT was designed to do more than survey the group being tested, to find the least common denominator. Teachers would like to see all of their students continue with their education to the extent of their interest and ability. We seriously doubt that anyone truly believes standards should be altered to fit the student insofar as measurement is concerned. Our research indicates that teachers use tests to determine where students are in their academic development so that educational plans may by formulated to meet the specific needs of the students. Exceptions to standards based on presumed differences are unfair and a disservice to students. It only masks and does not erase discrimination.

Slack and Porter continue with their criticism of the test makers, taking their concern to the limits of the debate. After scolding the Educational Testing Service (ETS) for requiring students to participate in validity studies against their will, without their knowledge, and without "rebate", the authors suggest another sinister possibility. The "Declaration of Helsinki", an agreement as to standards in clinical research, may have been violated because students are not told that they are involved in clinical research and have the right to withdraw at any time. Of course, this argument rests upon a rather broad and unusual definition of "clinical research." A rather unpleasant scenario emerges, upon considering the charge, of students unwittingly submitting themselves to the research only to wind up years later, writhing on the ground in the throes of multiple choice flashbacks.

It is not to say that nothing positive emerges from Slack and Porter's

study. Teachers about to be laid off due to enrollment declines can take comfort in the knowledge that medical school professors are insisting that anyone, given the opportunity and motivation, can become a doctor or lawyer or whatever he wants. High school grades, provided by teachers, turn out to be a better predictor of college achievement than the SAT according to the researchers. The only test missing as far as the doctors can tell seems to be a test of integrity. They seem to yearn for such a test, but stop short of describing how such an instrument might serve to predict success in college or assist in developing educational opportunities. The bad news is that the public will come away believing that the quality of education is not being measured, and it is less than important for success in college of later life. If you have recently had a graduate level course in measurement we suggest you read the study. If not, then wait a while as we are certain this is the kind of work that will be picked up and summarized in a future issue of Reader's Digest.

Another interesting piece of research was reported by Professor Ernest Sternglass, professor of radiological physics at the University of Pittsburgh. He recently told the participants at the American Psychological Association's annual meeting that the fallout from nuclear bomb tests was responsible for the decline in test scores for students born in Utah in 1958. While this study apparently was rather narrow, it says some very important things. Think of the school resources that would be used to educate these students without much hope of improving the students' achievement. It is a dangerous assumption to suggest that the geographical location of a student at his birth, or other environmental factors, is capable of determining lifetime achievement levels. Fans of Christopher Jencks may want to pursue this study to its obvious conclusion. The case it made to place less effort into teaching certain disadvantaged children because they are incapable of achievement.

Barbara Lerner, who has done considerable work on the legal aspects of testing, tells of the judge who presided in the Florida minimum competency case of Debra P. vs. Turlington. Federal District Court Judge George C. Carr ordered the State of Florida to issue regular diplomas to all public school students who attended high school for four years, regardless of whether or not they could read or write or do simple arithmetic. Not only did Judge Carr find that black students had been discriminated against because of the disproportinate number of black students who failed the test, he also found that students have a constitutionally protected right to a high school diploma which attests to their mastery of skills and curricula. Dr. Lerner draws upon the example of the physically ill person. "Would any responsible person assume that we could solve their problems by giving them all certificates of health?" posits Lerner. Again, the notion emerges that tests are the ends to be achieved, not part of a program of skill acquisition and education.

One of the more interesting points in the case was the penalty assessed against minorities who had to ride the buses for integration purposes. It was determined that the bus environment, owing to the length of time riding and the years involved, had an adverse effect on the students participating. This may appear as welcome news to groups attempting to limit bussing, and a handy argument for the next school board meeting. At the same time it poses a problem for educators in rural districts where children have been, are, and will continue riding buses to school and occasionally taking tests. If it is of any help, Judge Carr indicates lengthy rides correlate strongly with lower test scores. If your district has a lot of bussing and low test scores, you have another straw man.

Also, the all pervasive National Education Association has something to say about testing. They say that we should stop standardized testing, that those test result may be saying something significant about how kids achieve. What

the NEA cannot support empirically they seem to seek to achieve emotionally. By now you may have seen the slick newspaper ads depicting Melanie, the bright and happy first grader, who is not so happy anymore. In fact, they suggest the academically Melanie is damaged goods. The ad will say "...she has just been branded 'below average' by the people who make tests." She has been permanently damaged by taking a test. Somehow the NEA does not believe that teachers and schooling can make a difference in Melanie's life. No mention is made of teacher-made tests that some children fail to pass. Perhaps these children do not assume they are failures as a result of scores on those tests.

In the NEA advertisement you learn that students are subjected to "...many other tests...that don't measure what a student knows" and that parents need to write to their local NEA affiliate or the NEA in Washington to find out more about how their child is being categorized. The unspoken message in the ad is that nothing that teachers, administrators, or schools can offer in the remaining twelve years of public school can overcome the devastation to Melanie that occurred when she was labeled "below average." Whether or not NEA President Willard McGuire's minions will shell out the cash to have these camera-ready ads printed in your local paper remains to be seen. If and when it happens, though, there is something you can rely on besides the good sense of your parents and community members. You can count on the American Federation of Teachers to be out there supporting the proper and responsible use of all kinds of tests. We are going to be talking to teachers in districts throughout the country about the uses of standardized tests, ways of interpreting test scores, new tests for diagnostic and prescriptive purposes, and how to assist the media in understanding test scores.

In a national survey of elementary and secondary teachers who are members of the American Federation of Teachers, which was conducted in October 1979, it was found that teachers attached great importance to the use of standardized tests

for the diagnosis of students' needs and for student placement and grouping. They indicated that they used tests less so for determining class needs and judging individual student progress. In the perception of teachers, standardized tests were least useful for modification of course content and evaluation of instructional programs.

In spite of this very positive attitude toward standardized tests, teachers identified a variety of factors which inhibited their use of these tests. Major inhibitors of standardized tests are:

- Test results do not provide enough prescriptive information (63 percent of survey respondents)
- Test results do not provide an adequate profile of student strengths and wesknesses (52 percent)
- Test results are returned too late to be useful, or are not returned to teachers (51 percent)
- Test content does not match the curriculum (47 percent)
- Test materials are inappropriate for students (46 percent)

Teachers want to use standardized tests, they do find them useful, but have identified areas for change which would make them even more useful.

Teachers' attitudes and experiences with minimum competency testing were also surveyed. When asked their general attitude toward minimum competency testing, teachers responded as follows:

Favorable attitude	41 percent
Neutral attitude	38 percent
Unfavorable attitude	21 percent

Teachers generally seem positive in their attitude toward competency testing. Discussions with teachers on this subject indicate that teachers are interested in maintaining quality standards in schools and view competency testing as one method of doing that.

Teachers did have a number of specific observations and criticisms about minimum competency testing programs which now exist. These include:

- Teachers were not involved in developing competencies
- Teachers were not involved in setting passing scores
- Teachers receive little or no assistance in interpreting the test results or in planning remedial work

Teachers are rather mixed in their opinions about whether:

- important curricular goals have been ignored due to emphasis on minimum competency testing
- instructional programs have changed because of minimum competency testing
- minimum competency tests will promote higher quality instruction
- competencies to be tested reflect actual classroom goals

The survey suggests that teachers are generally supportive of the idea of minimum competency tests, but are quite critical of the way the idea is being implemented.

Similar surveys and studies have been conducted lately using different groups and instruments. These studies come after a decade which had seen a steady attack on tests and declining scores. The conclusions reached are similar to ours. One such study, by Michael Beck and Frank Stetz of the Psychological Corporation, surveyed a national sample of 3300 K-12 teachers about various standardized achievement test issues. They found the same generally positive attitudes toward most typical uses of standardized achievement tests that the AFT survey found.

Beck and Stetz consistently found two thirds of the teachers felt the amount of testing in their school sytem "about right." So much for the NEA's notion about too much testing. About 10 percent of the teachers personally used achievement test results "considerably" and about half made "some" use of the test results.

Given eight choices for test use, three-fourths of the teachers indicated frequent use was for diagnosing strengths and weaknesses, two-thirds said "measuring growth," and just under two-thirds said for "individual student evaluation."

As to how useful tests were for specific purposes, over three-fourths considered tests useful for measuring educational "growth" and detecting system-wide general strengths and weaknesses. Between 60 and 70 percent considered tests useful in reporting to parents, helping to plan instruction for class groups, helping to plan instruction for individual students, and for measuring the educational status of individual students. Teachers did not support the use of tests for reporting to newspapers, helping to evaluate teacher performance, and comparing classes within a school.

One of the findings of the study that supports AFT's position on tests was that fewer than one-sixth of the teacher favored the NEA position of moratorium against the use of all standardized tests. About one-fourth favored a moratorium on intelligence tests, while nearly one-third favored a moratorium on state mandated competency tests.

Beck and Stetz concluded that both teachers and students feel tests are generally useful. Teachers don't depend on tests completely, but neither do they want to give more of them. Students viewed tests generally as fair.

A recent recommendation by the Executive Council of the American Federation of Teachers is consistent with the findings of these surveys. Summarizing and clarifying resolutions that reach back into the early 1970's, this current statement supports the accurate and appropriate assessment of educational progress. Diagnosis of student problems can lead to appropriate prescriptive measures. The responsibility of public schools to show evidence of their proper functioning is judged important because of the relationship between public

perception of how well schools are meeting a broad range of educational needs and the support the public is willing to provide to education. The information from tests, adequately matched to the curriculum taught, can provide parents, teachers, and administrators with the kind of direction that is useful in planning for program changes and additions. Standards, and how they are being met, are a critical part of our defense of public education. If tests disappear, so do the standards they measure. We become vulnerable to the foes of public schools who will allege a greater decline in standards than exists. Without tests and standards, there is no defense.

AFT is not blind to the shortcomings of tests. Changes in existing programs and practices are desirable so long as the changes result in more accurate and useful information for educational purposes. Serious misuses of test data by public officials, errors and distortions in the media reporting of test data, and conclusions by the public and their representatives as to what a set of scores can mean must all be corrected. When comparisons between individual students are made based on test scores without reference to other factors which affect that performance, a disservice is done. Change, prudent change is welcome and useful. Sweeping changes such as those which would compromise the usefulness of the tests will be opposed. Testing regulation, commonly called "truth-in-testing" is such a change that is not consistent with our goals.

To implement these recommendations, and to address the problems outlined in this statement, AFT sought and received a grant from the National Institute of Education to assist teachers with the proper use of standardized tests. By the end of this year we will have completed regional workshops throughout the country.

The project (NIE grant No. NIE-G-79-0041) is to assist teachers in improving their use and understanding of standardized tests. One of the reasons that AFT became interested in testing initially was because there was some concern about the use of test data to evaluate and compensate teachers. The over reliance on tests was potentially damaging, especially when the use of tests was for purposes other than for which the test was intended. After surveying our membership we prepared a training manual to meet the interests and concerns of teachers. With assistance from the Center for the Study of Evaluation, UCLA, a manual of approximately 150 pages was developed. The manual describes standardized tests, how to select the tests to suit the teacher's purposes, how to interpret the results and scores of tests, and discusses the classroom applications. Some of the uses and abuses of tests are covered in the final chapter, and an appendix includes a review of the commonly used tests, a glossary, and bibliography.

We present the manual as part of a three-day training session in conjunction with discussions and presentations about some of the policy issues in testing. We have included speakers on truth in testing such as Senator Kenneth Lavalle, author of the New York truth-in-testing legislation and Robert Moulthrop of the Educational Testing Service. Dr. Christopher Pipho of the Education Commission of the States has contributed to the conferences by speaking on minimum competency testing. Assessment directors and administrators who are working in the field of testing at state and district levels have appeared to discuss their particular programs and the success or problems that they have been experiencing. Nationally known testing specialist Barbara Lerner presented the legal issues in testing and covered the wide range of topics from minimum competency to some of the implications that accompany truth in testing legislation at one conference.

The final draft of the manual will be available for participants to use in similar workshops in their states and local meetings. We will include in the final draft a summary chapter about the policy issues in testing. AFT will continue to update this section after the grant and provide the information to teachers periodically. It is important that teachers, who are involved in the administration of tests and the utilization of test data, have some idea of the kinds of issues in the debate that is developing. In the remaining training sessions we will include sections on teacher competency testing and alternative assessment techniques and instruments.

The conferences are scheduled on weekends to allow the greatest participation by classroom teachers. We begin our sessions on Friday evening with a conference overview and discussion of the issues. Manuals are available at this time for study. The second day of the conference is primarily devoted to the issues in the manual with sessions on policy issues in testing. In ten hours we are able to cover much of the information in the manual with satisfactory comprehension. The final day of the conference continues with some of the policy issues with practical exercises and discussions of testing concerns. While the program is quite rigorous, we find that the response is quite good and as a practical matter, the only alternative available. Most districts do not provide release time to their employees, even for conferences of this nature, and as a result, we must work with the teachers when we can.

Selection for attendance at the conference is made by the state federation office in each state. We ask that the president of each federation consider sending individuals who will be able to assist in future training sessions and workshops in their state. We provide through the grant the expenses of

the participants for food, travel, and housing. We have found this to be of great help in attracting participants from the organizations who are classroom teachers, not simply union staff or officers who might be able to find alternate financing for the meeting. In addition to the participants who we will fund, additional space is available for those who are interested in attending at their own expense. This group has included legislators, union officers and staff, school officials and representatives of other groups interested in education and testing. The sessions on policy issues have been the best attended by those outside the funded participants of the conference.

One of the most interesting developments from these conferences has been the concern with the way the media is reporting test data and the improper comparisons that are made between schools and districts by those within and outside of public education. Our participants indicate they will be seeking to enlist the aid of the local media in presenting test data in an appropriate and positive manner. The feeling seems to be that properly reported and analyzed, test data will be useful in program improvements to meet the concerns of students and parents. We have not found wide support for burying the results of tests, but rather teachers feel they have been asking for improvements that when provided will help in improving instruction. There is a confidence that the results of tests can be turned into a positive argument for support for the kinds of issues that frequently turn up at the bargaining table and in committee discussions at the school and district levels.

Another area that seems to be gaining support from the participants is a review of the curriculum with the goal being to match the tests given more closely with the curriculum actually taught. We have had inquiries as to methods for checking the match and what expectations for test scores can be

held when only a partial match is present. We have not heard demands for ways to teach to the test, but rather, a call for interpretation of test results when items tested have only partially been covered. I am certain those with responsibility for curriculum development at district and state levels have searched for the proper assessment tools, and yet the delivery of the information contained in the curriculum may not have the precise match with the available tests. In keeping with our belief that tests should not control the education process, but rather be one portion, AFT is exploring ways to help teachers evaluate the content of their course and test items available. Self assessment of content coverage can ease the work of district assessment directors, especially if ways are available to recover and utilize the information obtained. The close cooperation of teachers and their representatives will be the key to such a program.

The American Federation of Teachers feels that our efforts to assist in the use and understanding of this process will result in improvements that all may share.

We support the proper and appropriate use of standardized tests for diagnostic, prescriptive, selection, placement and program evaluation purposes, but we do not believe that testing information should comprise the sole basis for decision-making. Further, the AFT supports increased study of testing and the wide dissemination of information regarding the proper and appropriate use of standardized tests.

One overarching principle of the AFT position is that public education remains accountable to the public for its actions and the tests provide one, but not the only measure of educational performance.

To improve tests is to improve education. To destroy tests is to strike a dagger at the heart of educational quality and standards.